12th International Conference on Information Fusion
Seattle, WA, USA, July 6-9, 2009

System prediction combining state estimation with an evidential influence diagram

Johan Schubert, Pontus Svenson, Christian Mårtenson
Division of Information Systems
Swedish Defence Research Agency
SE-164 90 Stockholm, Sweden
{johan.schubert, pontus.svenson, christian.martenson}@foi.se
http://www.foi.se/fusion/

Abstract - *In this paper we develop a system state estimation model for combining partial information regarding the state of a system of interest. In addition we develop an evidential influence diagram representing our a priori knowledge of system relations. Both system state estimation and a priori knowledge are represented by belief functions. A predicted future system state is obtained by combining the fused estimated system state with the fused a priori knowledge. The predicted system state can be marginalized to give specific state predictions of all variables of interest of the system state estimation model. Finally, we may compare predicted system states with later actual states to highlight any deviations from expected developments.*

Keywords: System state prediction, system state estimation, influence diagram, situation assessment, threat assessment, Dempster-Shafer theory, belief function, effects-based approach to operations, knowledge support.

1 Introduction

In this paper we develop a system state estimation (SSE) method and an evidential influence diagram. Partial information regarding the state of a system may be combined using Dempster-Shafer theory [1], [2]. *A priori* knowledge (PK) regarding relations between system entities is represented by the evidential influence diagram. This knowledge is precombined and assumed unchanged during the system state estimation. Updating SSE by PK at each system state, we receive a system state prediction (SSP). The SSP may be marginalized for any system variable of interest yielding a specific belief function over the possible states of the variable of interest.

As an example of a system of interest, let us consider a system with several actors and some phenomenon. These actors may influence each other through multiple relations between subsets of their respective possible system states. Although phenomena do not act they may have the same type of influences between themselves, and both from and towards the actors, Figure 1.

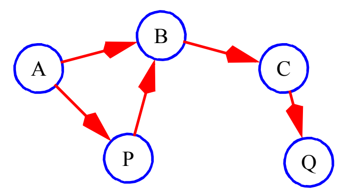

Figure 1. Influence diagram of actors and phenomena, i.e., A, B and C are actors and P and Q are phenomena.

Since both actors and phenomenon described have the same mathematical representation of their states and relations we will henceforth speak of entities and their influences.

We assume a reasoning chain where the system states are estimated sequentially throughout the analysis. At each system update we also perform a prediction of future system states by using the *a priori* knowledge of an evidential influence diagram. In Figure 2 a chain of reasoning is illustrated.

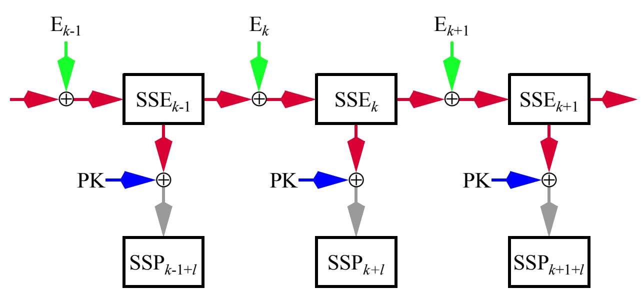

Figure 2. A chain of reasoning where system states are continuously updated as new evidence arrives and future system states are predicted after each update. E_k is the evidence becoming available at step k ($k \in \mathbb{Z}^*$), SSE_k is the system state estimation at step k, PK is the *a priori* knowledge and SSP_{k+l} is the system state prediction of step $k + l$, where l is the prediction length ($l \in \mathbb{R}$).

In Sec. 2 we develop a system state estimation model taking input data regarding states of single or multiple actors and phenomena. In Sec. 3 we introduce an evidential influence model describing our *a priori* knowledge of influence between actors and phenomena. Based on the estimated system state, we predict future system states using the influence model, Sec. 4. In Sec. 5 we perform predictions of future system states and compare these with the resulting system state at a later time. This can serve as a warning bell of an unexpected chain of events and as an assessment of the accuracy of the knowledge model.

2 System state estimation

2.1 Representation

Let us assume we have actors and phenomena. Each of these entities has an independent frame of discernment Θ_A. While our interpretation may differ they have the same type of representation. The cross product of all entities make-up the frame of discernment for the entire system,

$$\Theta = \times\{\Theta_A\} . \quad (1)$$

We further assume that each entity may have a multiple variable state. We have

$$\Theta_A = \times\{\Theta_{A_i}\}_i \quad (2)$$

where Θ_{A_i} is a set of alternative states for entity A_i. We have

$$\Theta_{A_i} = \{a_i^j\}_j . \quad (3)$$

2.2 Input data

Let input data in the general case be represented by belief functions on Θ, Eq. (1). We have

$$m_{E_k}^{\Theta}: 2^{\Theta} \to [0, 1] \quad (4)$$

where E_k is the kth piece of evidence.

We will represent a supported focal element χ_j of $m_{E_k}^{\Theta}$ on Θ as

$$\left\{(\theta_1, \ldots, \theta_{|\Theta|}) \middle| (\theta_1, \ldots, \theta_{|\Theta|}) \in \chi_j\right\}, \quad (5)$$

where $\chi_j \subseteq \Theta$. Note, that the sequence length is $|\Theta|$ (i.e., the number of entities). For an example, see Sec. 2.4.

The input intelligence is in the general case represented by basic belief assignments as

$$\left\{\left\{m_{E_k}^{\Theta}(\chi_j)\right\}_j\right\}_{k=1}^{|\Theta|} \quad (6)$$

where χ_j is the jth focal element and $\chi_j \subseteq \Theta$.

A special application interest of ours is input data in the form of dichotomous belief functions [2]. Recall that a dichotomous belief function is a belief function with only three focal elements χ, χ^c and Θ. This includes the case when input data is produced by Impactorium [3], a system for event recognition using Bayesian belief networks and soft/hard fusion developed at the Swedish Defence Research Agency. With data coming from Impactorium, we would have the additional requirement

$$m_{E_k}^{\Theta}(\Theta) \equiv 0 . \quad (7)$$

With a dichotomous belief functions we represent $m_{E_k}^{\Theta}$ on Θ as a basic belief assignment of

$$\left\{(\theta_1, \ldots, \theta_{|\Theta|}) \middle| (\theta_1, \ldots, \theta_{|\Theta|}) \in \chi_j\right\} \quad (8)$$

where $\chi_j \in \{\chi, \chi^c, \Theta\}$ and χ and χ^c are partitions of Θ, i.e., $\chi, \chi^c \subseteq \Theta$ where $1 \le |\chi| < |\Theta|$.

Thus, input data to the system state estimation process is represented as basic belief assignments

$$\left\{\left\{m_{E_k}^{\Theta}(\chi), m_{E_k}^{\Theta}(\chi^c), m_{E_k}^{\Theta}(\Theta)\right\}\right\}_k , \quad (9)$$

where χ and χ^c are focal element and a partition of Θ; $\chi, \chi^c \subseteq \Theta$ where $1 \le |\chi| < |\Theta|$.

2.3 Belief propagation

There has been some work on generally applicable improvements of the time complexity of Dempster's rule, e.g., [4], reducing time complexity in the general case from $O(2^{2|\Theta|})$ to $O(|\Theta|2^{|\Theta|})$. This is achieved by first using the fast Möbius transform on the basic belief assignments, transforming basic belief assignments into commonality functions. We have

$$Q_{E_k}^{\Theta}(\chi_i) = \sum_{\Theta \supseteq \chi_j \supseteq \chi_i} m_{E_k}^{\Theta}(\chi_j) . \quad (10)$$

This transformation is performed by function mtoq in TBMLAB [5]. Secondly, the commonality functions are then combined using Dempster's rule for commonalities (i.e., normalized multiplication of commonalities). We have

$$Q_{\{E_k\}}^{\Theta}(\chi_i) = \mathrm{K}\prod_k Q_{E_k}^{\Theta}(\chi_i) \quad (11)$$

where $\chi_i \neq \varnothing$ and K is a normalizing constant

$$\mathrm{K}^{-1} = \sum_{\varnothing \neq \chi_i \subseteq \Theta} (-1)^{|\chi_i|+1} \prod_k Q_{E_k}(\chi_i) . \quad (12)$$

Finally, from commonalities we transform back to basic belief assignments using

$$m_{\oplus\{E_k\}}^{\Theta}(\chi_j) = \sum_{\Theta \supseteq \chi_i \supseteq \chi_j} (-1)^{|\chi_i - \chi_j|} Q_{\oplus\{E_k\}}^{\Theta}(\chi_i) \quad (13)$$

which may also be performed using TBMLAB.

However, most improvements have concerned important special cases. Foremost among these are methods dealing with belief propagation in trees.

Depending on the structure of the input data, different algorithms for belief propagation can be used. These algorithms have different computational time complexities. In this section we review two different computational approaches for consideration.

The first algorithm discussed was developed by Shafer and Logan [6] for the case when evidence supports singletons or disjoint subsets of the frame. This is when a hierarchical network of subsets could be pruned to a hierarchical tree. The assumption is that a strict hierarchy of hypotheses can be defined from some subsets of 2^{Θ} and that a system will only receive information for these subsets[1]. The algorithm can handle evidence and calculate belief in partitions of the form $\{\chi_i, \chi_i^c\}$ for all subsets, χ_i, in the tree. It can also calculate belief in partitions of the form $C_{\chi_i} \cup \{\chi_i^c\}$, where C_{χ_i} is the set of children of χ_i.

The algorithm by Shafer and Logan can briefly be described as follows. The first step is borrowed from Barnett [7]. All evidence with equal foci, confirming and disconfirming, are combined, with the only difference that what Barnett did with simple support functions focused on singletons is done here for all subsets of the frame that are in the tree.

The two resulting simple support functions for each node are then combined into a belief function with focal elements $\{\chi_i, \chi_i^c, \Theta\}$.

Then we propagate the belief in the tree:

- For each parent, χ_i, of terminal nodes we combine the belief functions for all the children, store the belief in χ_i and χ_i^c at this parent node. This can be done with Barnett's method. These stored values will be used later when we propagate belief downwards through the tree. We combine the resulting belief function of the previous combination with the belief function for this parent and store the beliefs in χ_i and χ_i^c from the children and the parent at the parent node. These values will be used when we continue propagating belief upwards the tree. The procedure is repeated for the parents of these parents and so on climbing up the tree until we reach and store these values for the children of Θ.
- At the top we combine the belief functions from all the children of Θ and take one step down the tree to calculate the total belief in χ_i and χ_i^c for all children χ_i of Θ.
- Finally, we climb down the tree step-by-step until we reach all terminal nodes and calculate along the way the total belief in χ_i and χ_i^c for all subsets χ_i in the tree.

  First, in the last step down the tree we calculated the total belief in χ_i and χ_i^c for a particular parent. When climbing up the tree we stored with the same parent node the belief in χ_i and χ_i^c from the combination of all belief functions below the parent. From these values we can calculate the belief in χ_i and χ_i^c from all belief functions that are not below the parent.

  Secondly, for the parent, χ_i, we once again combine all belief functions that are below the parent, as we did when we climbed up the tree, but this time we find the belief in χ_j, χ_j^c and $A_j \cup A_i^c$, where χ_j is a child of χ_i. This is again done with Barnett's method.

  Thirdly, the total belief in χ_j and χ_j^c for a child is now found by combining these two belief functions from the first and the second step, i.e., the belief function of all subsets below the parent with the belief function of all subsets not below the parent.

[1]With Impactorium this is domain dependent as different dichotomous belief functions may intersect.

This algorithm has a linear time complexity in the number of the nonterminal nodes due to local computations and is linear in the tree's branching factor due to Barnett's approach.

If the conditions of the first algorithm are not fulfilled we use a second algorithm developed by Shafer, Shenoy and Mellouli [8] that propagates general belief functions in qualitative Markov trees. Qualitative Markov trees can arise through constructing a tree of families and dichotomies. This is done by substituting each nonterminal node with subset χ_i in a hierarchical tree by a parent-child pair with the dichotomy $\{\chi_i, \chi_i^c\}$ as subset at the parent, and the family $C_{\chi_i} \cup \{\chi_i^c\}$ as subset at the child and furthermore substituting terminal nodes with subset χ_i with the dichotomy $\{\chi_i, \chi_i^c\}$.

The algorithm for computing the total belief for every node in the tree of families and dichotomies is very simple:

- We will propagate belief to every neighbor, i.e., parent or child, χ_j of every node χ_i in the tree. We begin with the terminal nodes. Project the belief functions stored at every terminal node to its parent. For all neighbors but χ_j, if belief functions have been projected to χ_i from these neighbors then combine these belief functions and the belief function stored at χ_i and project the result towards χ_j. Whether or not belief has been projected from χ_j to χ_i is without significance to this rule.
- Finally, for every node χ_i in the tree, when belief functions have been projected from all neighbors χ_j we calculate the total belief in χ_i by combining all these projected belief functions and the belief function stored at χ_i.

This computational scheme reduces the time complexity from being exponential in |Θ| to being exponential in the size of the largest partition of Θ.

Resulting from these algorithms is a fused system state estimate

$$\begin{aligned} m_{\mathrm{SSE}_k} &= m^{\Theta}_{E_k} \oplus m_{\mathrm{SSE}_{k-1}}, \\ m_{\mathrm{SSE}_0} &= m^{\Theta}_{E_0}(\Theta) \end{aligned} \tag{14}$$

or

$$m_{\mathrm{SSE}_k} = \oplus \left\{ m^{\Theta}_{E_k} \right\}_k, \tag{15}$$

where the available evidence, Eq. (9) or Eq. (6), is combined by the first or second algorithm, respectively, Figure 2.

If an excessive amount of conflict builds up over time, this may be managed by discounting each piece of evidence in proportion to the degree that it contributes to the conflict [9] using the degree of falsity [10].

2.4 An example

Let us assume a problem with two entities, the weather forecast *W* and the risk of a riot *R*. In Sec. 3.4 and Sec. 4.3 we will describe how we can construct a causal relation between the two entities, but for now we just state that we have an interest in keeping track of them. W consists of one variable W_1 with three alternative propositions: *Sunny (Su)*, *Cloudy (Cl)* and *Rain (Ra)*. R also consists of one variable, R_1, which can take two different values, *Riot (Ri)* and *No riot (NRi)*. We have

$$\Theta = \Theta_W \times \Theta_R \tag{16}$$

where

$$\Theta_W = \Theta_{W_1}, \quad \Theta_R = \Theta_{R_1} \tag{17}$$

with

$$\Theta_{W_1} = \{Su, Cl, Ra\}, \quad \Theta_{R_1} = \{Ri, NRi\}. \tag{18}$$

Using Eq. (18) and Eq. (17) in Eq. (16) we get

$$\Theta = \Theta_{W_1} \times \Theta_{R_1} = \{(Su, Ri), (Su, NRi), \ldots, (Ra, NRi)\} \tag{19}$$

where $|\Theta| = 6$.

In general, input data take the form of basic belief assignments that may support any focal element of the frame of discernment Θ. However, a particular application interest are dichotomous belief functions with focal elements $\{A, A^c, \Theta\}$ where $A \subseteq \Theta$.

In the general case, input data take the form of Eq. (6) where the χ_j's are any subset of Θ in Eq. (19). In the case of dichotomous belief functions, input take the form of Eq. (9) where the χ's are any subset of Θ in Eq. (19) and the corresponding χ^c's are their complements $\Theta - \chi$ in Eq. (19).

We instantiate some dichotomous belief functions and study their combination. First we define our initial (*a priori*) system state estimation SSE_0. In the example we assume that we neither have information on the current mood of the population nor on the upcoming weather. Hence, we define

$$m^{\Theta}_{SSE_0}(\Theta) = 1.0. \tag{20}$$

At a later time we receive a weather forecast that tells us that there is a small chance that it will be sunny tomorrow and a somewhat larger chance that it will be cloudy. We denote the forecast as evidence E_1 and attach the following belief function

$$\begin{cases} m^{\Theta}_{E_1}(\{(Su,Ri), (Su,NRi)\}) = 0.3 \\ m^{\Theta}_{E_1}(\{(Cl,Ri), (Cl,NRi)\}) = 0.5 \\ m^{\Theta}_{E_1}(\Theta) = 0.2 \end{cases}. \tag{21}$$

In this particular case a new system state estimate SSE_1, will be identical to E_1.

3 Evidential influence model

3.1 Representation

The relations in our evidential influence diagram are between subsets of two entities of a system model. In Figure 3 a relation exist between subsets α_A and β_B of entities A and B.

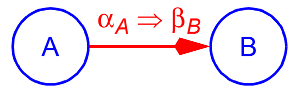

Figure 3. A two entities influence diagram.

The implication $\alpha_A \Rightarrow \beta_B$ for any entities A and B is a proposition with a certain basic belief assignment, where $\alpha_A \subseteq \Theta_A$ and $\beta_B \subseteq \Theta_B$ are subsets of their respective frames of discernment. Note, that α_A and β_B may be subsets of the possible states any individual variables A_i and B_j, $\alpha_A \subseteq \Theta_{A_i}$ and $\beta_B \subseteq \Theta_{B_j}$, or any cross product of the possible states of several A_i's and B_j's including Θ_A and Θ_B, Eq. (2). We will however always represent the *a priori* knowledge as subsets of the entire frame Θ, Eq. (1).

The implication $\alpha_A \Rightarrow \beta_B$ can be stated in set theoretic terms as

$$\alpha_A^c \cup \beta_B \tag{22}$$

and is formally a subset of the frame of discernment $\Theta = \times\{\Theta_A\}$, Eq. (1), where $\Theta_A = \times\{\Theta_{A_i}\}_i$, Eq. (2).

We will thus represent the basic belief assignment of $\alpha_A^c \cup \beta_B$ as a function on Θ.

We have α_A and β_B represented on Θ as

$$\alpha_A: \quad \left\{(\theta_1, \ldots, \theta_A, \ldots, \theta_{|\Theta|}) \middle| \theta_A \in \alpha_A, \forall j \neq A.\ \theta_j \in \Theta_j\right\} \tag{23}$$

and

$$\beta_B: \quad \left\{(\theta_1, \ldots, \theta_B, \ldots, \theta_{|\Theta|}) \middle| \theta_B \in \beta_B, \forall j \neq B.\ \theta_j \in \Theta_j\right\}, \tag{24}$$

respectively, where $\alpha_A \subseteq \Theta_A$ and $\beta_B \subseteq \Theta_B$.

The cardinality of the sets in Eq. (23) and Eq. (24) are

$$|\alpha_A| \prod_{j \neq A} |\Theta_j| \tag{25}$$

and

$$|\beta_B| \prod_{j \neq B} |\Theta_j|, \tag{26}$$

respectively.

Furthermore, α^c is represented on Θ as

$$\alpha_A^c: \quad \left\{(\theta_1, \ldots, \theta_A, \ldots, \theta_{|\Theta|}) \middle| \theta_A \notin \alpha_A, \forall j \neq A.\ \theta_j \in \Theta_j\right\} \tag{27}$$

with cardinality

$$|\alpha_A^c| \prod_{j \neq A} |\Theta_j|. \tag{28}$$

The union $\alpha_A^c \cup \beta_B$ is a focal element of $m_{(A,B)}^{\Theta}$ which must be written as union of Eq. (27) and Eq. (24) with cardinality less or equal to the sum of the cardinalities of Eq. (28) and Eq. (26). It can be rewritten in different ways, but not simplified to a single set of sequences.

3.2 Instantiating *a priori* knowledge

Let

$$m_{(A,B),r}^{\Theta}: 2^{\Theta} \to [0, 1] \tag{29}$$

be the rth basic belief assignment on the entire frame Θ expressing a relation between A and B, where all our *a priori* knowledge visualized in the influence diagram is represented as

$$\left\{ \left\{ m_{(A,B),r}^{\Theta}(\alpha_A^{ic} \cup \beta_B^i) \right\}_i, m_{(A,B),r}^{\Theta}(\Theta) \right\}_r \tag{30}$$

where i is the ith focal element of basic belief assignment $m_{(A,B),r}$ and $\alpha_A^{ic} \cup \beta_B^i$ may be substituted by the union of Eq. (27) and Eq. (24), and Θ may be substituted by Eq. (5) or Eq. (8), respectively.

3.3 Combination of *a priori* knowledge

These basic belief assignments can be pre-fused as we assume the *a priori* knowledge to be unchanged during the system state estimation and prediction process.

We have *a priori* knowledge

$$m_{PK} = \oplus \left\{ m_{(A,B),r} \right\}_r . \tag{31}$$

As we assume that the *a priori* knowledge is unchanged during the system state estimation process the computational time complexity is not an issue here.

The basic belief number $m_{PK}(\Theta)$ may be viewed as the inertia of the system during update with the *a priori* knowledge. When $m_{PK}(\Theta)$ is high then the prediction SSP_{k+l} is little changed from the estimate SSE_k.

3.4 The example

Let us return to the example previously discussed. We assume a problem with two entities, W and R, with one variable each and three (*Su*, *Cl* and *Ra*) and two (*Ri* and *NRi*) alternative propositions, respectively. The frame of discernment is given by Eq. (19).

We now want to extend our example model by adding *a priori* knowledge of how the entities influence each other. The knowledge is elicited from two Subject Matter Experts (SMEs), SME_1 and SME_2. SME_1 claims that cloudy or rainy weather will reduce the risk of a riot. SME_2 on the other hand claims that sunny or cloudy weather will increase the risk of a riot. The *a priori* knowledge of the system will be the combination of these statements.

For the *a priori* knowledge from SME_1, α_W^c instantiates as

$$\alpha_W^c = \{(Su, Ri), (Su, NRi)\} \tag{32}$$

and β_R instantiates as

$$\beta_R = \{(Su, NRi), (Cl, NRi), (Ra, NRi)\} \tag{33}$$

where

$$\alpha_W^c \cup \beta_R = \{(Su, Ri), (Su, NRi), (Cl, NRi), (Ra, NRi)\} . \tag{34}$$

The *a priori* knowledge from SME_2, α_W^c instantiates as

$$\alpha_W^c = \{(Ra, Ri), (Ra, NRi)\} \tag{35}$$

and β_R instantiates as

$$\beta_R = \{(Su, Ri), (Cl, Ri), (Ra, Ri)\} \tag{36}$$

where

$$\alpha_W^c \cup \beta_R = \{(Su, Ri), (Cl, Ri), (Ra, Ri), (Ra, NRi)\} . \tag{37}$$

In Eq. (30) we have a general description of *a priori* evidence. Here, Eq. (34) and Θ are the two focal elements of the basic belief assignment of SME_1 where

$$\begin{cases} m_{(W,R)}^{\Theta}(\{(Su, Ri), (Su, NRi), (Cl, NRi), (Ra, NRi)\}) = 0.4 \\ m_{(W,R)}^{\Theta}(\Theta) = 0.6 \end{cases} \tag{38}$$

and Eq. (37) and Θ are the two focal elements of the basic belief assignment of SME_2 where

$$\begin{cases} m_{(W,R)}^{\Theta}(\{(Su, Ri), (Cl, Ri), (Ra, Ri), (Ra, NRi)\}) = 0.3 \\ m_{(W,R)}^{\Theta}(\Theta) = 0.7 \end{cases} . \tag{39}$$

The basic belief numbers of each basic belief assignment given by the SMEs indicate their confidence in these statements.

Combining the two basic belief assignments yields

$$\begin{cases} m_{(W,R)}^{\Theta}(\{(Su, Ri), (Ra, NRi)\}) = 0.12 \\ m_{(W,R)}^{\Theta}(\{(Su, Ri), (Su, NRi), (Cl, NRi), (Ra, NRi)\}) = 0.28 \\ m_{(W,R)}^{\Theta}(\{(Su, Ri), (Cl, Ri), (Ra, Ri), (Ra, NRi)\}) = 0.18 \\ m_{(W,R)}^{\Theta}(\Theta) = 0.42 \end{cases} . \tag{40}$$

4 System state prediction

Predicting future states of a system is a very general problem in information fusion as well as in many other research areas. Perhaps the easiest to understand example of system state prediction is the tracking problem. In this problem, we are given uncertain observations of an objects position and are asked to, at any time, determine its current position. The solution of this problem is to maintain an estimate of the current position which is updated whenever new observations arrive. If both the time between observations and the uncertainty of the observations are small, we can take the last observation as the current prediction. If these conditions are not met, however, we must perform an estimation of the current prediction. Usually, this is done in two steps. First, a motion model is used to update the current estimate of the position. Then, this estimate is updated by fusing it with any new observations which might have arrived. The motion model attempts to model the system dynamics and corresponds to the *a priori* knowledge given in Eq. (31). If the motion model is accurate enough, it is in principle possible to predict the state of the system at any arbitrary time in the future.

4.1 Representation

Having derived the representations of the system state estimation SSE in Eq. (5) or Eq. (8) (depending on the structure of input data) and the representation of the *a priori* knowledge PK of the evidential influence model as the union of Eq. (27) and Eq. (24), we may find the representation of their combination. The representation of the focal elements of the resulting prediction from combining the latest estimate SSE_k with the *a priori* knowledge PK is simple. For a particular focal element, we have

$$\left\{ \bigcap \left\{ (\theta_1, \ldots, \theta_{|\Theta|}) \mid (\theta_1, \ldots, \theta_{|\Theta|}) \in \chi_j^s \right\}_{s \in S} \right\} \cap \left\{ \bigcap \left\{ \alpha_A^{ic} \cup \beta_B^i \right\}_{i \in I} \right\} \quad (41)$$

where χ_j^s is the jth focal element of $m_{E_s}^{\Theta}$ and $\alpha_A^{ic} \cup \beta_B^i$ is the ith focal element of $m_{(A,B),r}^{\Theta}$, with $S \subseteq K$, the index set of $\{m_{E_k}^{\Theta}\}$, and $I \subseteq R$, the index set of $\{m_{(A,B),r}^{\Theta}\}$.

This is the representation of the system state prediction received from an earlier system state estimation using the *a priori* knowledge PK. We have

$$m_{\mathrm{SSP}_{k+l}} = m_{\mathrm{PK}_{k+l,k}} \oplus m_{\mathrm{SSE}_k}, \quad (42)$$

see Figure 2.

4.1.1 Multiple predictions

We may consider using several prediction steps without intermediate update, Figure 4.

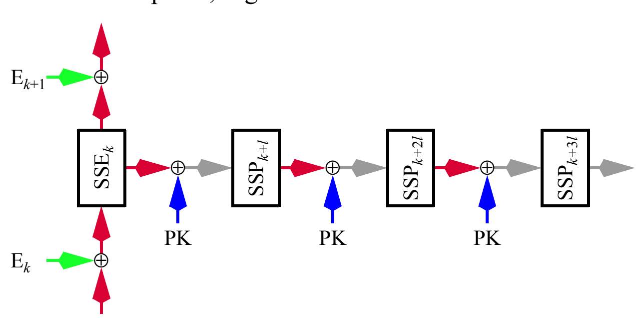

Figure 4. Multiple prediction steps SSP_{k+il} ($i \in \mathbb{Z}^+$).

This is appropriate from a statistical point of view. The *a priori* knowledge fused with the system state estimation and successive system state predictions is not combined with itself. In fact, each application of a prior knowledge is information to predict system changes at different system steps. We have

$$\begin{aligned} m_{\mathrm{SSP}_{k+(i+1)l}} &= m_{\mathrm{PK}_{k+(i+1)l,\,k+il}} \oplus m_{\mathrm{SSP}_{k+il}},\ i \geq 1 \\ m_{\mathrm{SSP}_{k+l}} &= m_{\mathrm{PK}_{k+l,\,k}} \oplus m_{\mathrm{SSE}_k} \end{aligned} \quad (43)$$

or

$$m_{\mathrm{SSP}_{k+(i+1)l}} = \oplus \left\{ m_{\mathrm{PK}_{k+(j+1)l,\,k+il}} \right\}_{j=0}^{i} \oplus m_{\mathrm{SSE}_k} \quad (44)$$

where $\forall ij.\ \mathrm{PK}_i = \mathrm{PK}_j$, Figure 4.

However, we must make sure that the successive system state predictions remain within a tolerable margin of error. This can be monitored by comparing predicted and actual system states when the system has evolved to the predicted step. Note, that this comparison involves a measure of interpolation as predicted system states at steps $k + il$ ($k \in \mathbb{Z}^*$; $i \in \mathbb{Z}^+$; $l \in \mathbb{R}$) are not synchronized with the sequence of estimated system states at steps k, as l belongs to the reals. This is studied in Sec. 5.

4.2 Specific states

For every SSE_k and SSP_{k+il} we may marginalize the distribution to any entity of interest. Marginalization is a way of focusing a belief function on subset of Θ. The marginalizations that are of interest to us are marginalizations on the frame Θ to any individual Θ_A. We have

$$m_{\mathrm{SSP}_{k+l}}^{\downarrow\Theta_A}(X) = \sum_{Y^{\downarrow\Theta_A} = X} m_{\mathrm{SSP}_{k+l}}^{\Theta}(Y) \quad (45)$$

for all Θ_A and k, and identically for m_{SSE}.

4.3 The example

Resulting from combining the latest estimate SSE_1 with the *a priori* knowledge PK we receive a prediction SSP_{1+l}. Combining the two corresponding belief functions yields

$$\left\{ \begin{aligned} m_{SSP_{1+l}}^{\Theta}(\{(Su, Ri)\}) &= 0.16 \\ m_{SSP_{1+l}}^{\Theta}(\{(Cl, Ri)\}) &= 0.06 \\ m_{SSP_{1+l}}^{\Theta}(\{(Cl, NRi)\}) &= 0.09 \\ m_{SSP_{1+l}}^{\Theta}(\{(Su, Ri), (Su, NRi)\}) &= 0.36 \\ m_{SSP_{1+l}}^{\Theta}(\{(Su, Ri), (Ra, NRi)\}) &= 0.02. \\ m_{SSP_{1+l}}^{\Theta}(\{(Cl, Ri), (Cl, NRi)\}) &= 0.13 \\ m_{SSP_{1+l}}^{\Theta}(\{(Su, Ri), (Su, NRi), (Cl, NRi), (Ra, NRi)\}) &= 0.06 \\ m_{SSP_{1+l}}^{\Theta}(\{(Su, Ri), (Cl, Ri), (Ra, Ri), (Ra, NRi)\}) &= 0.04 \\ m_{SSP_{1+l}}^{\Theta}(\Theta) &= 0.08 \end{aligned} \right. \quad (46)$$

Our concern in this example scenario is whether we should expect a riot or not. This is found by marginalizing the basic belief assignment to Θ_R. We have

$$\begin{cases} m_{SSP_{1+l}}^{\downarrow\Theta_R}(Ri) = 0.22 \\ m_{SSP_{1+l}}^{\downarrow\Theta_R}(NRi) = 0.09 \\ m_{SSP_{1+l}}^{\downarrow\Theta_R}(\Theta) = 0.69 \end{cases} . \quad (47)$$

Thus, we have some indication of an upcoming riot while the uncertainty is large.

5 Comparing predicted and actual system states

As outlined in Sec. 4, the system state prediction output will be updated whenever the system state estimation (Sec. 2) has changed. In practice, this will be done as soon as there is enough new input data suggesting that the system state estimation has changed.

Depending on the temporal scale used when constructing the influence model described in Sec. 3, it will be necessary to update the system state estimation, Eq. (14), more or less often. For some system state prediction applications, the prediction step is followed by an updating step, where new system state observations are fused with the predicted system state estimate and the result is used to get a new predicted system state. Perhaps the most familiar example of such a system is ordinary tracking. In such systems, the information processing algorithm can generally be written as

- given a system state estimate for step k and a motion model that describes how the state changes with time,
- predict system state at step $k + 1$ by applying the motion model to the system state estimate at step k,
- fuse this predicted system state with observations collected at step $k + 1$ to produce a new system state estimate for $k + 1$.

This procedure is then applied recursively to obtain a system state estimate for all times. There are two crucial assumptions made in this procedure. First, it must be possible to compute a motion model that allows us to predict the single-step behavior of the system. This is often possible for simple systems that move simply, such as ships on the surface of the sea or aircraft [11], but considerably more difficult for more complicated motions [12] or systems. The procedure also breaks down if the time between observations of the system state is too large.

In our case, we do not attempt to track the system state continuously using the influence model in Eq. (31) as a motion model. Instead, we apply Eq. (42) as soon as we have new data, and it is this system state estimate that is shown to the user in the decision support. Our reason for doing this is simple: since we are dealing with considerable more complicated problems than tracking of aircraft, we believe it to be naïve to expect a “motion model” to be able to give more than single-step predictions of the system state. Instead, we use observation data as a situation assessment at all times, and only use the prediction/influence model to produce a situation prognosis.

However, doing a prediction on the predicted system state is still useful for other purposes.

Applying the method introduced in [13], we can perform comparisons of the system prognosis at some step k with evolved versions of earlier prognoses as illustrated in Figure 5. If there is a significant discrepancy between the different prognoses, the system can alert the user to this fact using a “red signal flag” or “alarm bell” that informs the user that something strange has happened. Getting such warnings would be extremely useful if, for instance, the user has done their planning based on the earlier prognosis and we can now tell, using more recent observation data, that the other actors are doing something different from what we expected. It could also be used to detect possible attempt at deception from opposing actors.

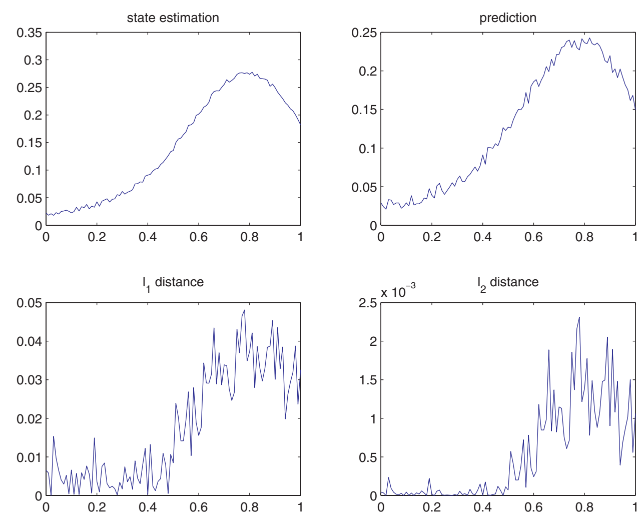

Figure 5. Comparison between system state estimation and prediction using the L_1 and L_2 norms.

Calculating the discrepancy between the prognoses at different steps could be done in several different ways [13], [14]. Perhaps the most straightforward is to simply calculate the L_p-distance

$$L_p(m_1, m_2) = \left(\sum_{x \subseteq \Theta} |m_1(x) - m_2(x)|^p \right)^{\frac{1}{p}} \quad (48)$$

between the mass functions of the different system state prognoses. A slightly more sophisticated approach would to calculate the Kullback-Leibler divergence

$$D_{KL}(m_1, m_2) = \sum_{x \subseteq \Theta} m_1(x) \log \frac{m_1(x)}{m_2(x)} . \quad (49)$$

In addition, there are several more advanced measures that could be used. Any of the similarity measures commonly used to compare probability distributions could be used to compare the mass functions of interest.

The system state estimates and prognoses calculated are probabilistic, since both the data that is fed into it and the influence model used are stochastic. This means that care must be taken when displaying the results to the user: should the visualization display the complete mass function or some most-likely outcome from it (e.g., BetP [15])?

In Figure 5, we show the results of comparing a system state estimation and a system state prediction using the norms introduced in Eq. (48) with $p = 1$ and $p = 2$. For this visualization, we chose to project the mass functions onto the real line and not attempt a display of the values of the mass function for all elements of the lattice 2^{Θ}.

6 Discussion

A possible generalization of the method described in this paper is to include entity copies for all steps k. Thus,

$$\Theta = \times\{\Theta_{A,k}\}_{A,k} \tag{50}$$

where $\Theta_{A,k}$ is the set of possible system states on entity A at step k and Θ_k is the set of possible system states at step k.

With this representation we may represent *a priori* knowledge regarding relations between any entity at step k and the same entity at step $k+1$, e.g.,

$$\beta^c_{B,k} \cup \beta_{B,k+1} . \tag{51}$$

This is a more specific modeling of system state changes of entity B from k to $k+1$ than the inertia discussed in Sec. 3.3.

The method presented in this paper can be understood as a generalization of the concept of dynamic Bayesian networks [16]. In dynamic Bayesian networks, one considers distinct but identically-structured networks for each step, but also allows some of the nodes in the network at step k to be connected to nodes in the network at step $k + 1$. Conceptually, this is similar to Figure 2.

7 Conclusions

We have developed an evidential system state estimation and system state prediction model. These models represent observations of system states as belief functions and *a priori* knowledge regarding all possible state transitions as one fused belief function.

With these models we can perform system state predictions of the entire system as well as individual state predictions of all system entities through marginalization of the prediction belief function.

References

[1] A. P. Dempster, "A generalization of Bayesian inference," *Journal of the Royal Statistical Society Series B,* Vol. 30, No. 2, pp. 205−247, 1968.

[2] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.

[3] P. Svenson, T. Berg, P. Hörling, M. Malm, and C. Mårtenson, "Using the impact matrix for predictive situational awareness," in *Proceedings of the Tenth International Conference on Information Fusion*, Québec, Canada, 9−12 July 2007, Paper 1298, pp. 1−7.

[4] R. Kennes, and P. Smets, "Computational aspects of the Möbius transformation," in *Uncertainty in Artificial Intelligence 6*, Machine Intelligence and Pattern Recognition Series, Vol. 12, P. P. Bonissone, M. Henrion, L. N. Kanal and J. F. Lemmer (Eds.), Elsevier Science Publishers B.V. (North-Holland), Amsterdam, 1991, pp. 401−416.

[5] P. Smets, (2005, Mar.), "TBMLAB 3.1," [Online], Available: http://iridia.ulb.ac.be/~psmets/#G

[6] G. Shafer, and R. Logan, "Implementing Dempster's rule for hierarchical evidence," *Artificial Intelligence*, Vol. 33, No. 3, pp. 271−298, 1987.

[7] J. A. Barnett, "Computational methods for a mathematical theory of evidence," in *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, Vancouver, Canada, 24−28 August 1981, pp. 868−875.

[8] G. Shafer, P. P. Shenoy, and K. Mellouli, "Propagating belief functions in qualitative markov trees," *International Journal of Approximate Reasoning*, Vol. 1, No. 4, pp. 349−400, October 1987.

[9] J. Schubert, "Conflict management in Dempster-Shafer theory by sequential discounting using the degree of falsity," in *Proceedings of the Twelfth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, L. Magdalena, M. Ojeda-Aciego, J. L. Verdegay (Eds.), Málaga, Spain, 22−27 June 2008, pp. 298−305.

[10] J. Schubert, "Specifying nonspecific evidence," *International Journal of Intelligent Systems*, Vol. 11, No. 8, pp. 525−563, August 1996.

[11] S. Blackman, and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House Publishers, Norwood, MA, 1999.

[12] H. Sidenbladh, "Multi-target particle filtering for the probability hypothesis density," in *Proceedings of the Sixth International Conference on Information Fusion*, Cairns, Australia, 8−11 July 2003, pp. 800−806.

[13] H. Sidenbladh, P. Svenson, and J. Schubert, "Comparing future situation pictures," in *Proceedings of the Eighth International Conference on Information Fusion*, Philadelphia, USA, 25−29 July 2005, Paper C10-4, pp. 1−6.

[14] H. Sidenbladh, P. Svenson, and J. Schubert, "Comparing multi-target trackers on different force unit levels," in Proceedings of SPIE Vol. 5429 *Signal Processing, Sensor Fusion, and Target Recognition XIII*, I. Kadar (Ed.), Orlando, USA, 12−14 April 2004, pp. 306−314.

[15] P. Smets, and R. Kennes, "The transferable belief model," *Artificial Intelligence*, Vol. 66, No. 2, pp. 191−234, April 1994.

[16] K. P. Murphy, Dynamic Bayesian networks: representation, inference and learning, Ph.D. Thesis, University of California, Berkeley, CA, 2002.